AD-A116 897 RAND CORP SANTA MONICA CA F/G 12/2

ARE ROBUSTNESS MEASURES ROBUST?(U)

JAN 82 R Y ARGUEDEN

UNCLASSIFIED RAND/P-6734 NL

1 OF 1

AD A116897

END

DATE FILMED

8-82

DTIC

ARE ROBUSTNESS MEASURES ROBUST?

R. Yılmaz Argüden

January 1982

P-6734

ARE ROBUSTNESS MEASURES ROBUST?*

R. Yılmaz Argüden

The Rand Corporation

January 1982

---

*The author, R. Yılmaz Argüden is a graduate fellow in the Rand Graduate Institute. This paper was prepared for an RGI tutorial, "Robustness in Operations Research," given by Dr. James Bigelow in the Fall Quarter 1981.

ABSTRACT

To a large degree, the classical approach to problem solving in operations research (OR) is to fit a real life situation into a well-known OR model. When OR models are used to deal with major policy problems in which the underlying processes are not well understood, this effort results in too much simplification.

Due to an inability to perceive all uncertainties, and a consequent wish to retain flexibility once the decisions are made, decisionmakers are more interested in the "robustness" of their policy decisions than their "optimality," which becomes a vague concept due to the nature of these problems.

This paper emphasizes the desirability of robustness and criticizes attempts to fit an operationalized measure of robustness into an optimization structure, by the aid of a decision analytic example.

I. INTRODUCTION

The role of operations research (OR) models in public-policy decisionmaking have been viewed in two philosophically different ways: (1) As tools to be utilized to obtain the optimal solution to a policy problem, and (2) as a means of providing insights into a problem area, illuminating the inherent trade-offs in alternative solutions, screening alternatives, and generating alternatives for further exploration. The virtues of the second view, to which I subscribe, have been described elsewhere (for example, [1] and [5]).

Some who have recognized the drawbacks of using models to obtain an "optimal" solution (some of the drawbacks are detailed below) have suggested using robustness as a remedy. However, in operationalizing the idea they have tried to quantify robustness. This paper criticizes attempts to calculate robustness values and argues for other operationalizations of the idea, in view of recent trends in OR.

II. SPECTRUM OF PROBLEMS

OR models have been applied to a wide range of problems since the 1950s. Although vastly different in nature they can be characterized by two groups which span the range, say Type A (well defined) and Type Z (fuzzy).

Type A Problems

These usually have a widely accepted single objective, or an objective which can serve as a good proxy for all the others. Alternatives are easy to identify. There is little room for unperceived or unquantified uncertainties, so risks can usually be incorporated into

a model using probabilities. Models built for these problems, more often than not, represent the real situation closely. Type A problems are generally tactical in nature. Decisions, once taken, are applied repeatedly to the situation. Hence, expected value can be used as the relevant measure of payoff. OR models have found wide application to this class of problems (e.g., determining input-output mix for refineries).

Type Z Problems

These problems have a multi-objective nature and generally the weights to be given to different objectives are highly subjective. Therefore, multiple objectives cannot be evaluated by a single, universally-accepted criterion. Some issues are qualitative and it may be very difficult to introduce them into the model in a quantitative manner. For this reason, many relevant issues may remain unmodelled. Uncertainties are extensive, both because of the nature of the problem and because of the unmodelled issues. Furthermore, some uncertainties cannot be expressed in terms of probabilities, mainly because they cannot be fully perceived ahead of time. Finally, Type Z problems involve decisions which are generally strategic in nature. Their reversibility is very limited and costly, and their consequences are long-standing. Therefore, decisionmakers are more interested in the robustness of the solutions than in their expected payoffs. Type Z problems can be characterized by large investment decisions and major public policy issues.

Application of OR models, and for that matter any model, to this type of problem has raised major concerns with decisionmakers and some policy analysts.

III. ROBUSTNESS TO THE RESCUE

Being aware of the difficulties of utilizing OR models to find an "optimal" solution to Type Z problems, some OR practitioners have suggested using robustness as a relative criterion

Robustness is a characteristic of an alternative. It may be defined as the insensitivity to the assumptions of a model, and to the value judgements placed on different objectives of a problem.

Therefore, alternatives which will perform well under many different states of nature and on most or all the dimensions under consideration are considered robust. Decisionmakers are increasingly more interested in "robust solutions" rather than "optimal" ones as they are involved more with problems near the Type Z end of the spectrum.

IV. ROBUSTNESS MEASURES AND THEIR DRAWBACKS

The recognition that solutions may be robust with respect to only some of the assumptions, illuminates the difficulty of coming up with a general measure of robustness. Nevertheless, a number of papers written on the subject have attempted to measure robustness (e.g., [3], [6], [7], [8], [9], [10], [12]).

In the following, I use one of the robustness measures mentioned in the literature as an example to illustrate some of the drawbacks of these measures, namely their sensitivity to the definition of "satisfactory outcomes," the number of alternatives considered, and unbalanced payoff values.

Rosenhead, et al. [7] view robustness as retaining flexibility in the initial moves of a multi-stage decisionmaking process. In their words, "A plan whose initial decisions limit the future as little as possible has an evolutionalry advantage in an uncertain world." They define a measure of robustness, r_1 for a decision, d_i as

$$r_i = \frac{n(\hat{s}_i)}{n(\hat{s})}$$

where $n(\hat{s})$ is the number of satisfactory outcomes available and $n(\hat{s}_i)$ is the number of satisfactory outcomes attainable after the decision, d_i. Let's apply this definition to the following example to show some of the weak points.

Example 2.

The California government is interested in whether or not to create incentives to promote crude oil production; to do so, \$5* will

---

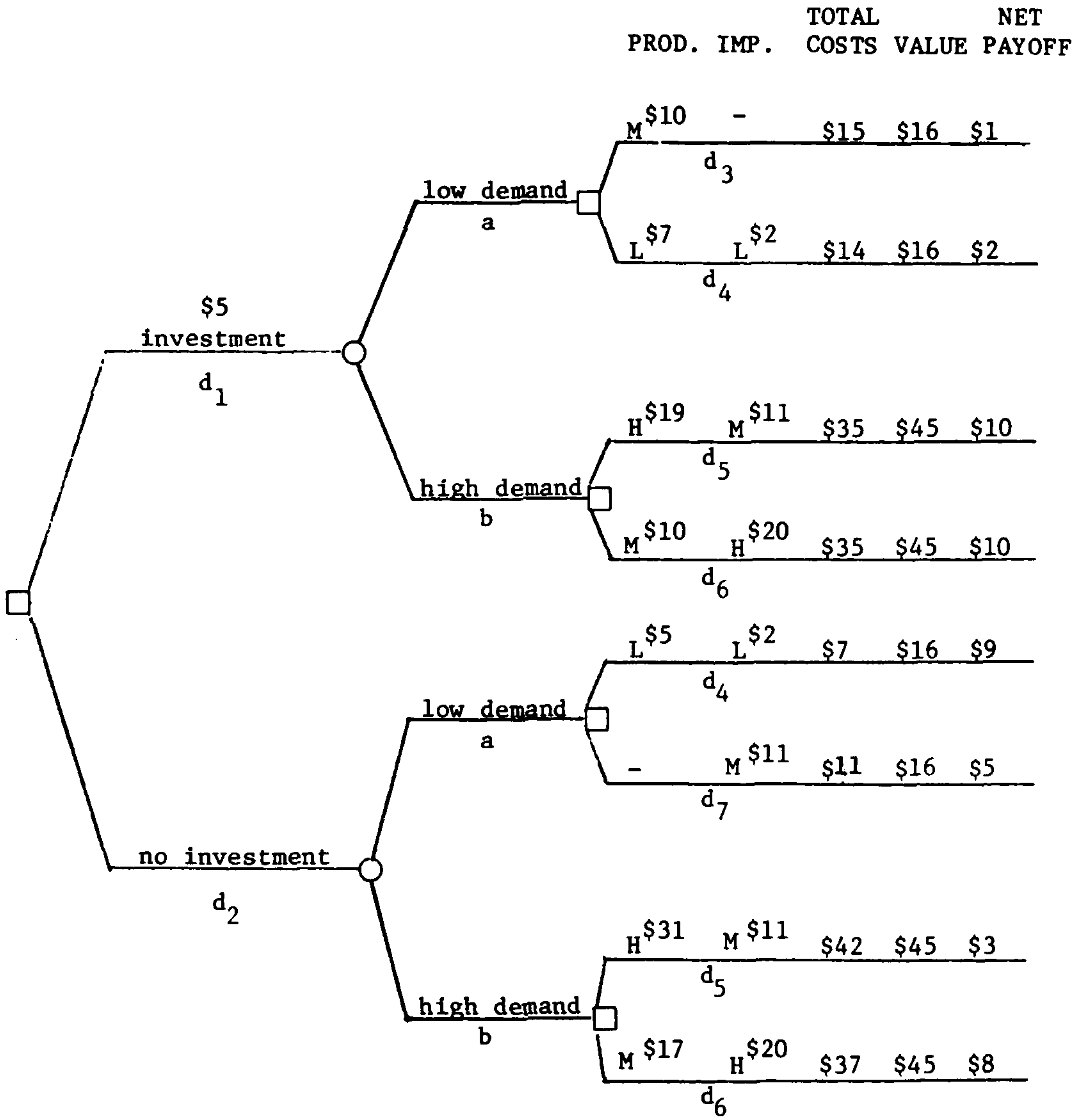
*All figures are in billions of dollars and represent present value of actual amounts that will be spent or earned in the next 15 years. The numbers are hypothetical.

have to be invested in equipment for a complex but more efficient oil production technology. If these investments are undertaken, production costs will decrease for medium and high production levels from $17 to $10 and from $31 to $19, respectively, but for low production levels operating costs will increase by $2 from the previous $5 level, due to the more complex technology. Crude oil demand will be met by a combination of in-state production and imports. The three levels of imports being considered: low, medium, and high; are expected to cost $2, $11, and $20, respectively. If demand turns out to be low, it can be met separately by a medium level of imports, or by medium production, or jointly by low import and low production. Crude oil in the case of low demand will have a value of $16.* If demand turns out to be high, crude oil will be valued at $45 and the demand can be met either with high imports and medium production or with medium imports and high production. a and b are the probabilities of low and high demand, respectively. This information can be displayed on a decision tree, omitting outcomes with negative payoffs.

---

*The value of crude oil is assumed to be unaffected by production decisions.

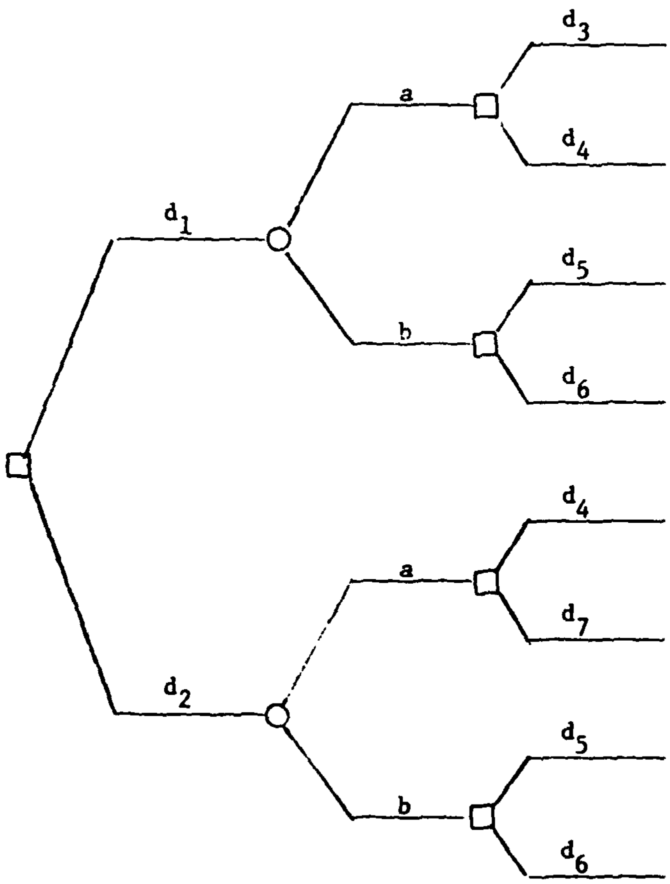
FIGURE I



□ denotes a decision node and O denotes a chance node.

Sensitivity to the Definition of "Satisfactory"

FIGURE IA (Same as Fig. I)



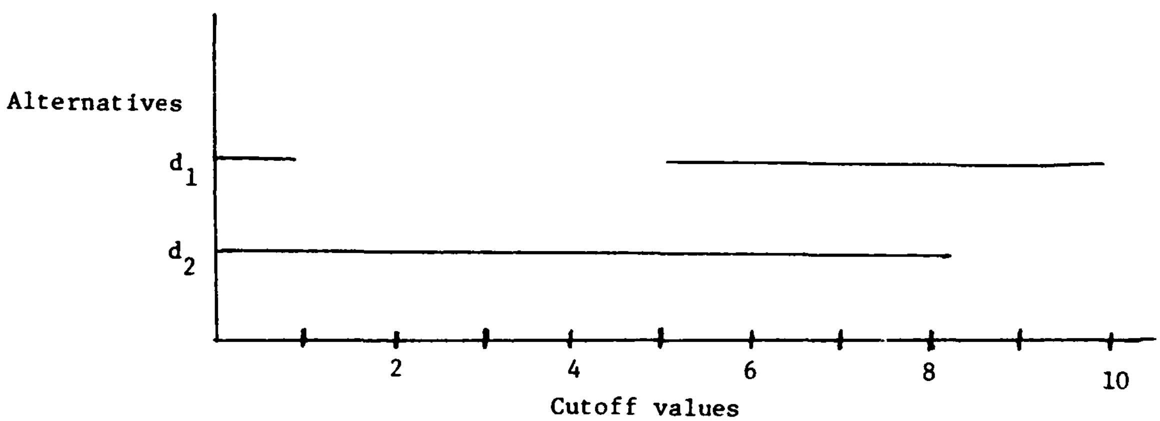
Payoff	Satisfactory (cutoff = 8)
1	0
2	0
10	1
10	1
9	1
5	0
3	0
8	1

Let a cutoff value be defined as the minimum payoff value which the decisionmakers will accept as satisfactory.

If the cutoff point for a satisfactory outcome is chosen to be eight, then $r_1 = r_2 = 2/4$, hence robustness of d_1 and d_2 are equal, but the choice of a cutoff point is arbitrary and it can influence

the measure in undesirable ways. For example, if the cutoff point was chosen to be nine, in Figure IA, r_1 would be greater than r_2 $(r_1 = 2/3 > 4_2 = 1/3)$. Graph I, which shows the preferred alternatives versus cutoff values, may be informative and may help for the choice of a cutoff point.

Graph I



If $1 < \text{cutoff} \leq 5$, d_2 is better than d_1 with respect to robustness. If $8 < \text{cutoff} \leq 10$, d_1 is better with respect to robustness. For other cutoff values, robustness does not discriminate between the alternatives in this example.

Sensitivity to Assumed Payoff Values

One way to get around the above problem would be to use the values

of the payoffs rather than the zero-one coding assigned to each outcome (by its magnitude w.r.t. the cutoff point). Such a measure will give the following robustness scores, rr_i for decisions in Ex. 2.

$$rr_1 = \frac{23}{48} \qquad rr_2 = \frac{25}{48}$$

Note that measure rr_i assumes equal probabilities for a and b (a reasonable assumption in case of absolute ignorance).

Another way may be the comparison of the lengths of the regions of cutoff points where one decision is preferred over the other. (See Graph I.) The measure, l_i , will give the following scores for Ex. 2.

$$l_1 = 10 - 8 = 2 \qquad l_2 = 5 - 1 = 4$$

This would imply that d_2 is preferable.

In still a different version, Pye [4] incorporates the ideas of values and satisfactory outcomes in his last measure of robustness by assigning a value of zero to unsatisfactory outcomes and using the actual payoff values for satisfactory outcomes.

The problem with using values is that a very large value will distort the measures. For example, in Ex. 2, if one of the payoffs, say 10, were 1000 instead, it would have changed the robustness measures mentioned in this section in favor of d_1 even if the probability of b were very low (i.e., 10^{-6}).

Also, when the payoffs are multi-attribute in nature, reducing them to a single value may be very hard and controversial, but

classifying the alternatives as satisfactory and unsatisfactory may be much easier.

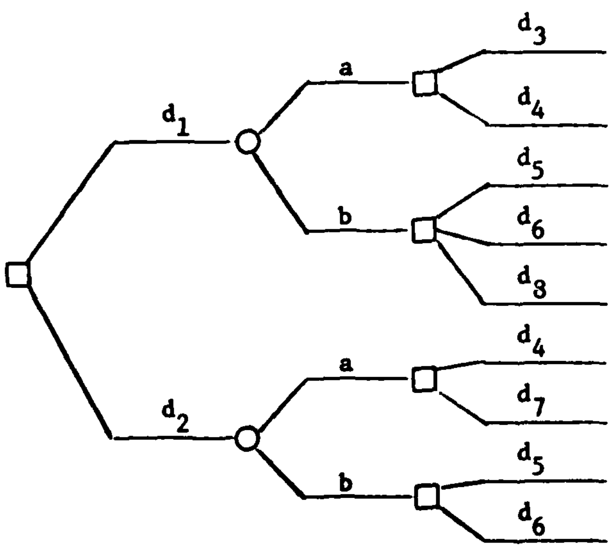
Insensitivity to the Location of Retained Flexibility

The next example is illustrative of the point that Rosenhead's measure r_i is insensitive to the location of the satisfactory outcomes after the decision between d_1 and d_2.

Example 2a.

This example is the same as the previous one, except that an alternative is added: If investments are undertaken, production levels may be increased to a very high level (v.h.), d_8, at a cost of \$31. Then the decision tree will look as follows:

FIGURE II



Payoff	Satisfactory (cutoff = 8)	
1	0	
2	0	
10	1	$r_1 = \frac{3}{5}$
10	1	
9	1	
9	1	
5	0	$r_2 = \frac{2}{5}$
3	0	
8	1	

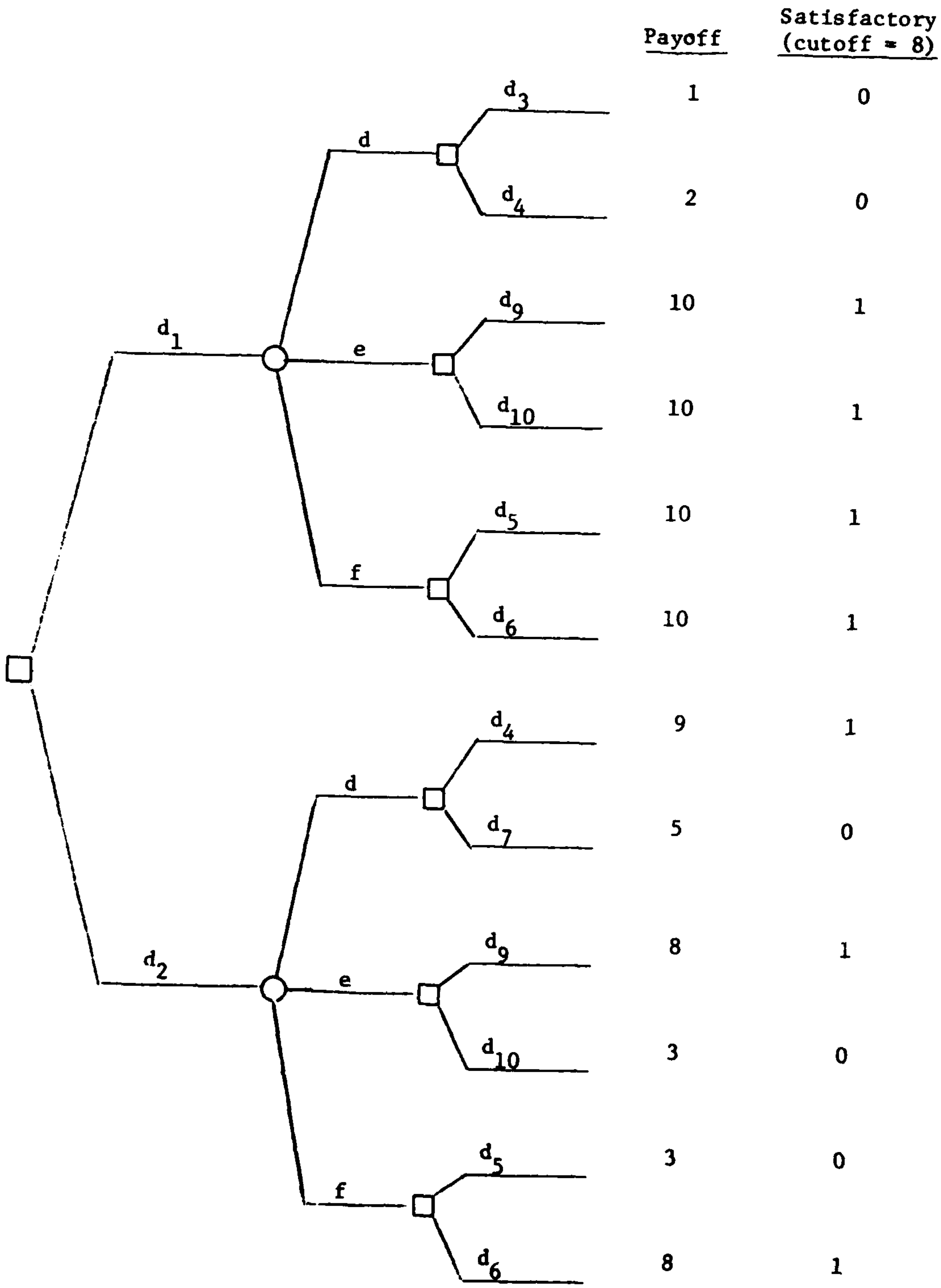
The measure does not account for the fact that all three of the satisfactory outcomes after d_1 lie on branch b, i.e., can be exercised only if nature picks b. In contrast, there is a satisfactory outcome no matter what nature picks if the initial decision were d_2. Therefore, r_i misleads us in terms of robustness.

Sensitivity to the Number of Alternatives Considered

Example 2 can be elaborated by including a medium demand level scenario. Let d, e, and f be the probabilities of high, medium, and low demand, respectively. Let's assume that medium demand can be met either with medium production and medium imports, d_9, or with high production and low imports, d_{10}. The value of crude oil will be $36 in this case. Therefore, if investments were undertaken, d_9 and d_{10} would cost (10 + 11 + 5=) $26 and (19 + 2 + 5=) $26. Otherwise, d_9 would cost (17 + 11=) $28 and d_{10} would cost (31 + 2=) $33. Figure III summarizes the new situation. When e + f = b, this example is equivalent to Ex. 2, but the measure r_i gives significantly different answers with these two different ways of grouping alternatives. When e + f is marginally different from b, which model to choose remains as a subjective decision to the analyst.

Therefore, the robustness measures mentioned in the literature themselves are not robust with respect to the definition of satisfactory outcomes, number of alternatives considered, and unbalanced payoff values. Clearly, measuring robustness is by definition a futile effort since it requires knowledge of unperceived uncertainties.

FIGURE III



$r_1 = \frac{4}{7}$ $\quad$ $r_2 = \frac{3}{7}$

Furthermore, using a formula to calculate robustness values and using these values for comparison purposes, is nothing more than optimizing with respect to a new criterion (i.e., "robustness" replaces "optimization"). Hence, it is open to the same criticisms which lead the way to searching robust solutions.

V. ARE WE OUT OF HOPE?

Inability to measure robustness does not imply inability to operationalize the idea. In fact, some recent trends, which originated from the viewpoint of using OR models as tools to gain insights rather than to find the "optimal solution", are towards operationalizing the idea of robustness.

For example, displaying information about each alternative on each dimension of the problem to clarify the inherent trade-offs (e.g., the "scorecard" method devised by Goeller et al. [4]; or [5]) is such an effort. This is valuable because it is very difficult and in certain cases practically impossible (i.e., when there is a large group of decisionmakers who cannot agree on the relative merits of different objectives) to reduce the dimensions under consideration into one. Also, communicating the insights gained through modelling effort to the decisionmakers by displaying information, will leave the decision to the decisionmakers and allow them to incorporate hidden, unmodelled issues into their decisions.[11] Where some issues cannot be modelled quantitatively, a qualitative analysis of those issues will be of more use to the decisionmakers than studies which ignore them.

Greater utilization of recent advances in computer technology by user-interactive computer modelling may enhance analysts' capabilities

to respond to decisionmakers' informational needs [13].

In a paper yet to be published, Brill, et al. [2] address the issue of devising methods to generate alternatives (the Hop-Skip-Jump method). Throughout this paper I have elaborated on the need for finding satisfactory robust solutions rather than optimal ones. Therefore, one needs to evaluate many alternatives which have satisfactory outcomes (if there isn't any, one probably has to reduce his expectations); Brill's paper deals with identifying different alternatives.

VI. CONCLUSIONS

OR models have been and will be useful to the decisionmakers. However, model building and alternative identification are both influenced by analysts' subjective decisions, and their choices may have major implications on the conclusions of an analysis. Constraints imposed by the way analysis is conducted, when not explicitly stated, may reduce the decisionmaker's confidence in the study, hence diminish the usefulness of the analysis as an aid to decisionmaking. Therefore, all limitations imposed on the problem by the modelling efforts should be made explicit.

In real life problems it is very hard to find dominating or even robust solutions with respect to all the assumptions and impacts. Even then alternatives that are robust with respect to only some of the assumptions may still be worth taking by the decisionmakers. Although robustness of an alternative is important to consider, it should not be used as the sole criterion to evaluate alternatives especially since robustness cannot be reduced to a number. Therefore, robustness of different alternatives should be identified, possibly by using sensitivity

analysis, and discussed. Analyzing many alternatives and communicating the insights gained to the decisionmakers will receive better acceptance than prescribed "optimal solutions" which are too sensitive to the assumptions to be applicable.

ACKNOWLEDGEMENTS

I wish to thank Jim Bigelow for the stimulating discussions which made this paper possible. Jan Chaiken's and Warren Walker's criticisms on an earlier draft are gratefully acknowledged. Gene Fisher, Charles wolf, Jr., and Syam Sarma also kindly provided valuable comments. Of course, the responsibility of the opinions and any remaining errors are solely the author's.

REFERENCES

1. Brill, E. D., "The Use of Optimization Models in Public-Sector Planning," Management Science, Vol. 25, No. 5, May 1979.

2. Brill, E. D., et al., "Modelling to Generate Alternatives: The HSJ Approach," forthcoming in Management Science.

3. Elton, M., et al., "Robustness and Optimality Criteria for Decisions--Reply," Operations Research Quarterly, Vol. 24, No. 2, 1973.

4. Goeller, B. F., et al., Protecting an Estuary from Floods: A Policy Analysis of the Oosterschelde, Vol. 1, Summary Report, The Rand Corporation, R-2121/1-NETH, December 1977.

5. Liebman, J. C., "Some Simple-Minded Observations on the Role of Optimization in Public Systems Decision Making," Interfaces, Vol. 6, No. 4, August 1976.

6. Pye, R., "Formal Decision Theoretic Approach to Flexibility and Robustness," Journal of the Operational Research Society, Vol. 29, No. 3, 1978.

7. Rosenhead, J., et al., "Robustness and Optimality as Criteria for Strategic Decisions," Operations Research Quarterly, Vol. 23, No. 4, 1972.

8. Rosenhead, J., "Education in Robustness," Journal of the Operational Research Society, Vol. 29, No. 2, 1978.

9. Rosenhead, J., "Note on Robustness and Interdependent Decision-Making," Journal of the Operational Research Society, Vol. 30, No. 6, 1979.

10. Schenkerman, S., "Robustness of Discrete Decisions to Worst Case Distribution," Journal of the Operational Research Society, Vol. 29, No. 2, 1978.

11. Schilling, D. A., A. McGarity, and C. ReVelle, "Hidden Attributes and the Display of Information in Multiobjective Analysis, Forthcoming in Management Science.

12. White, D., "Robustness and Optimality Criteria for Decisions," Operations Research Quarterly, Vol. 24, No. 2, 1973.

13. Walker, W. E., Models in the Policy Process: Past, Present, and Future, The Rand Corporation, P-6654, September 1981.

BIBLIOGRAPHY

Abrahamse, A. F., et al., Protecting an Estuary from Floods: A Policy Analysis of the Oosterschelde, Vol. II, Assessment of Security From Flooding, The Rand Corporation, R-2121/2-NETH, April 1977.

Anderson, A., "Application of Post Sample Stability Theory to Time Series Forecasting," Operations Research Quarterly, Vol. 28, No. 1, 1977.

Baumgartner, T., T. R. Burns, and L. D. Meeker, "The Description and Analysis of System Stability and Change: Multilevel Concepts and Methodology," Quality and Quantity, Vol. 11, 1977.

Bigelow, J.H., et al., Protecting an Estuary from Floods: A Policy Analysis of the Oosterschelde, Vol. III, Assessment of Long-Run Ecological Balances, The Rand Corporation, R-2121/3-NETH, April 1977.

Box, G. E. P., and N. R. Draper, "Robust Designs," Biometrica, Vol. 62, No. 2, 1975.

Gardner, Bruce, "Robust Stabilization Policies for International Commodity Agreements," American Economic Review, Vol. 69, May 1979.

Hogg, R. V., "An Introduction to Robust Procedures," Comm. in Statistics, Part A, 1977.

Ignall, E. J., et al., "Using Simulation to Develop and Validate Analytic Models: Some Case Studies," Operations Research, Vol. 26, No. 2, March-April 1978.

Infante, E. F. and J. L. Stein, "Optimal Growth with Robust Feedback Control," Review of Economic Studies, Vol. 40, January 1973.

Kramer, G. H., "Robustness of the Median Voter Result," Journal of Economic Theory, Vol. 19, 1978.

McLean, S. and T. Abodunde, "Entropy as a Measure of Stability in a Manpower System," Journal of the Operational Research Society, Vol. 29, No. 8, 1978.

Relles, D. A. and W. H. Rogers, "Statisticians are Fairly Robust Estimators of Location," Journal of American Statistical Association, Vol. 32, No. 357, Theory and Methods Section, March 1977.

Robinson, S. M., "Characterization of Stability in Linear Programming," Operations Research, Vol. 25, No. 3, 1977.

Stigler, S. M., "Do Robust Estimators Work with Real Data?," The Annals of Statistics, Vol. 5, No. 6, 1977.

Vassiliou, G., "A Note on Stability in Gani-Type Models in Manpower Systems," Journal of the Operational Research Society, Vol. 31, No. 10, 1980.